A Study on Significant Predictors for Prediction of Undiagnosed T2DM Using Binary Logistic Regression Model

S. S. N. Zainal1*, M. J. Masnan1, A. Ahmed1, N. A. M. Amin1 and M. I. Omar @Ye Htut2

1Institute of Engineering Mathematics, Universiti Malaysia Perlis, Pauh Puara Campus, 02600 Arau, Perlis, Malaysia
2University Health Centre, Universiti Malaysia Perlis, Pauh Puara Campus, 02600 Arau, Perlis, Malaysia.
*Corresponding author Email: sitilsababilahmobilah@yahoo.com

Abstract

Type 2 Diabetes Mellitus (T2DM) is a chronic disease that can cause premature deaths worldwide. Malaysia is one of the many countries that facing this serious epidemic. The World Health Organization (WHO) has also estimated that Malaysia would have 2.8 million people having T2DM disease in 2030. This study aims to identify significant predictors for prediction of undiagnosed T2DM patients in one of the highest prevalence states of T2DM. Binary logistic regression model proposed to predict the presence of T2DM among undiagnosed respondents. The selection of significant predictors using univariate, multivariate and backward stepwise selection was implemented in this study. The study concludes that four predictors were found significant for prediction of undiagnosed T2DM patients.

Keywords: Binary Logistic Regression model; Significant predictors; undiagnosed T2DM

1. Introduction

The prevalence of Type 2 Diabetes Mellitus (T2DM) is rapidly increasing and can affect the health of people worldwide [1][2]. According to [3], there are 2 major types of diabetes mellitus which are Type 1 diabetes mellitus (T1DM) and T2DM disease. T1DM disease is defined as insulin-dependent where the body does not enough produce insulin while T2DM is defined as non-insulin dependent where the body ineffectively uses insulin. Besides, diabetes mellitus disease can occur among the pregnancy woman that called as gestational diabetes mellitus (GDM) disease. GDM disease can be characterized as a temporary condition that occurs during pregnancy and can put at risk as T2DM disease for a long term. However, T2DM disease is one of the main health care problem that threaten to reach pandemic by 2030.

Latest finding from the National Health & Morbidity Survey (NHMS II) in 2015, the prevalence of total diabetes mellitus has risen to 17.5% where 8.3% of prevalence were among individual with known diabetes mellitus and 9.2% of prevalence among individual with undiagnosed diabetes mellitus. The prevalence of diabetes mellitus is referred as the percentage of the number of diabetes mellitus cases in a population. Figure 1 shows the trend of total prevalence for diabetes mellitus aged 18 years and above from the NHMS report for five alternate years [4][5][6]. Those report revealed that the prevalence of diabetes mellitus has increased at 11.6% in 2006 and 17.5% in 2015. Undiagnosed diabetes mellitus become alarming cases where it started to increase slowly at 0.8% difference from known diabetes mellitus in 2011 and continually increase in 2015. According to the newspaper report [7], 50% from 3.6 million Malaysians i.e. about 1.8 million people are believed to have the T2DM disease where they are not yet diagnosed and never have health screening procedure.

Fig. 1: Trend of total prevalence for diabetes mellitus aged 18 years and above [4][5][6].

In primary clinical care, binary logistic regression model has been applied to investigate related predictors associated with diabetes mellitus. There are twelve predictors have been identified by [8] in predicting the diabetes or prediabetes mellitus which include age, family history of diabetes, marital status, educational level, work stress, duration of sleep, physical activity, gender, eating fish, drinking coffee, preference for salty food, and body mass index (BMI). While, some significant predictors suggested by [9] might be indicated to the chances for having diabetes mellitus disease were age, BMI, hypertension, dyslipidemia, impaired fasting glucose and impaired glucose tolerance. Another researcher found that the significant predictor of BMI category (i.e. who are obese about 1.5 to 5 times higher than individuals have normal BMI) is strongly associated to the risk of T2DM disease [10]. Thus, this research aims to identify the significant predictors in predicting undiagnosed T2DM patients using binary logistic regression model.